INTERFERENCE DETECTION IN GNSS SIGNALS USING THE GAUSSIANITY CRITERION

Fernando D. Nunes*, Fernando M. G. Sousa†

* Instituto de Telecomunicações and IST
† Instituto de Telecomunicações and ISEL

ABSTRACT

We assess the performance of Gaussianity tests, namely the Anscombe-Glynn, Lilliefors, Cramér-von Mises, and Giannakis-Tsatsanis (G-T), with the purpose of detecting narrowband and wideband interference in GNSS signals. Simulations have shown that the G-T test outperforms the others being suitable as a benchmark for comparison with different types of interference detection algorithms.

Index Terms—Gaussianity tests, interference detection, kurtosis, cumulants

1. INTRODUCTION

Current Global Navigation Satellite Systems (GNSS) use spread-spectrum modulation schemes to carry navigation information but the transmitted signals present often poor robustness to certain types of RFI (radio frequency interference) which can be intentionally provoked (jamming) or due to the proximity of authorized radio sources, such as TV/FM, radar and mobile systems. These interfering systems tend to exhibit narrow bands compared with the GNSS signal bandwidths, whereas jammers typically use wideband signals for malicious interference [1]. Thus, we will consider two types of interfering signals: narrowband sinusoidal and wideband FM-modulated chirp signals.

In [2], a method for pre-correlation interference detection in GNSS signals was proposed. It was considered a semi-blind interference detection scheme as it compares the characteristics of the signal disturbed by RFI with an interference-free signal. Although being more sensitive to weak interference than the absolute detection schemes, it requires that the GNSS signal is sampled at a window (assessment window) where it is known that the signal is RFI-free (see also [3]). However, (i) often this assumption cannot be guaranteed or (ii) the interference detection algorithm has to operate in a scenario where RFI is already present. In those cases, it would be convenient to use blind detection schemes that perform reasonably well without data from an assessment window. This consideration is the starting point for the present work.

Typical sampled complex baseband received signals contain data from several satellites and constellations together with uncorrelated Gaussian samples due to thermal noise. As a consequence, the inphase/quadrature components are approximately zero-mean Gaussian provided that the number of RF signals impinging the receiver’s antenna at a given band is larger than about five. For a large number of RF signals Gaussianity is guaranteed by the Central Limit theorem [4]. In practical scenarios this condition is met most of the time; that is, with the deployment of several GNSSs besides GPS, the number of satellites in view is almost always larger than five [5].

The tests of Gaussianity (or normality) have been used in many different applications to separate (classify) two populations that obey to different statistical distributions when it is known that one of them is Gaussian [6]. There are a plethora of tests available which are usually based on the evaluation of the distribution moments such as the kurtosis and the skewness (e.g., Anscombe-Glynn and Jarque-Bera tests), or the maximum distance between two points of the experimental and reference curves of the distribution (e.g., Kolmogorov-Smirnov and Lilliefors tests), or the average distance between experimental and reference curves (Cramér-von Mises test). In [7], Giannakis and Tsatsanis proposed a test for Gaussianity that relies on higher-order statistics. Specifically, the algorithm resorts to third or fourth-order cumulants to measure deviations from Gaussianity and uses the fact that, if a sequence of real samples are Gaussian, then the k-th order cumulants vanish for \( k \geq 3 \). Although computationally heavier, the 4th-order cumulants are preferable as the 3rd-order cumulants (as well as any odd order above the first [8]) are zero for any symmetrical distribution and in most cases will fail to detect departure from Gaussianity. Therefore, we will consider herein only 4th-order cumulants.

This paper aims at analyzing and comparing the performance of normality tests in the detection of RFI in GNSS receivers. The tests to be considered are the Anscombe-Glynn [9], Lilliefors [10], Cramér-von Mises [11], and Giannakis-Tsatsanis [7]. The comparison will be carried out mainly through the receiver operating characteristic (ROC) plots which are curves of correct interference detection probability versus false alarm probability for a pre-defined scenario.

2. SIGNAL CHARACTERIZATION

Assume that the incoming signal disturbed by interference (jamming signal) \( s(t) \) is [5]

\[
r(t) = \sum_{g=1}^{N_s} A_g c_g(t) d_g(t) \cos(2\pi(f_0 + f_g)t + \phi_g) + s(t) + w(t)\]

(1)
where \( N_0 \) is the number of received GNSS signals and \( w(t) \) is AWGN with power spectral density \( G_w(f) = N_0/2 \). For each received signal, with amplitude \( A_g \), the carrier-to-noise ratio is \( (C/N_0)_g = A_g^2/(2N_0) \). \( c_g(t) \) is the spreading sequence, \( d_g(t) \) is the data sequence, \( f_0 \) and \( f_g \) are, respectively, the nominal carrier and Doppler frequencies, and \( \phi_g \) is the initial carrier phase. We consider two types of continuous interferers: a sinusoidal waveform of frequency \( f_0 + f_j \), with \( f_j \) accounting for the jammer’s frequency offset, and a chirp waveform, both described by \( s(t) = A_J \cos(\varphi(t)) \), where \( A_J \) stands for the interference amplitude. The jamming-to-signal power ratios are \( (J/C)_g \).

The sinusoidal (narrowband) interference is defined as
\[
s(t) = A_J \cos(2\pi f_J t + \varphi_J),
\]
with \( \tau = \Delta^{-1} \) to yield
\[
z_m = s_m + A_J \exp[j(2\pi f_J t_m + \varphi_J)] + n_m \tag{2}
\]
with \( t_m = m\Delta \) and \( n_m = N_{g,m} + jN_{q,m} \). The noise components \( N_{g,m} \) and \( N_{q,m} \) are independent Gaussian zero-mean random variables (r.v.) with variance \( N_0/\Delta \). Signal
\[
s_m = \sum_{g=1}^{N_g} A_g c_{g,m} d_{g,m} \exp[j(2\pi f_g t_m + \varphi_g)] \tag{3}
\]
is the contribution of the visible satellites. The pdfs of \( A_g c_{g,m} d_{g,m} \cos(2\pi f_g t_m + \varphi_g) \) and \( A_g c_{g,m} d_{g,m} \sin(2\pi f_g t_m + \varphi_g) \) are given by
\[
f_y(y) = 1/(\pi \sqrt{A_g^2 - y^2}), \quad |y| < A_g,
\]
with variance \( A_g^2/2 \).

For \( N_g \geq 5 \) independent GNSS signals with equal amplitudes \( A_g = A \), \( g = 1, \ldots, N_g \), the real and imaginary parts of \( s_m \) are well approximated by zero-mean Gaussian r.v. with variances \( A^2 N_g/2 \), thanks to the Central Limit Theorem. Therefore, in the absence of an interfering signal and assuming that several satellites are in view, we can consider that the real and imaginary parts of the samples \( z_m \) are independent and approximately zero-mean Gaussian with variances greater than \( N_0/\Delta \).

In [1] several civil jammers were tested and the signals found were, typically, of the chirp type with periods of repetition in the order of several microseconds and frequency sweeps exceeding 20 MHz. Based on the analysis carried out in [1] we model the wideband interference as a periodic chirp signal with the frequency sweep depicted in Fig. 1, where \( L \) stands for the frequency sweep range, and \( T_J \) is the period of the jamming signal.

The instantaneous frequency is given by \( f_i(t) = f_0 + \gamma t - L/2 \) in the interval \( 0 \leq t \leq T_J \), with \( \gamma = L/T_J \). Let \( \tau = \text{mod}(t, T_J) \) be the remaining after division of \( t \) by \( T_J \); the interfering signal is given by
\[
s(t) = A_J \cos(2\pi f_0 - L/2) + \pi \gamma^2 t \cos(\varphi_0).
\]

Signal \( r(t) \) with chimp interference is heterodyned to baseband and sampled at rate \( r \) to yield
\[
z_m = s_m + A_J \exp[j(\pi \gamma^2 t_m - \pi L t_m + \varphi_0)] + n_m \tag{4}
\]
with \( \tau = \text{mod}(t_m, T_J) \). The signals amplitudes are given by
\[
A_g = \sqrt{2N_0/(C/N_0)}_g = \sigma_N \sqrt{2\Delta(C/N_0)}_g \text{ and } A_J = A_g \sqrt{(J/C)_g} = \sigma_N \sqrt{2\Delta(C/N_0)}_g(J/C)_g.
\]

We consider sampling rates such that \( A_J^2/2 \ll \sigma_J^2 \).

### 3. Gaussianity Tests

In this section several Gaussianity tests will be described. In all cases hypothesis \( H_0 \) corresponds to complex Gaussian observations \( z_m \) and hypothesis \( H_1 \) corresponds to non-Gaussian observations. The former is considered to characterize a GNSS signal without interference and the latter is assigned to a signal affected by interference.

- **Anscome-Glynn test.** A normally distributed stationary sequence \( x(i), i = 1, \ldots, N \) or its random variable \( X \sim N(\mu, \sigma^2) \) may be characterized by its \( k \)-order central moments, which are zero, if \( k \) is odd, and \( 1 \times 3 \times \cdots (k-1) \beta^k \), if \( k \) is even [4]. Of particular interest are the third and fourth standard moments given by
\[
E(x^3) = E((X - \mu)^3)/\sigma^3 \text{ and } b_2 = E((X - \mu)^2)/\sigma^2 \text{ that measure skewness and kurtosis, respectively.}
\]

For the Gaussian distribution they are given by \( b_3 = 0 \) and \( b_2 = 3 \). The non-normality of a given sequence can be measured by the deviation of the skewness and kurtosis from the values taken by the Gaussian distribution. However, the skewness is not useful if the non-normal distribution remains symmetric. Therefore, for certain types of signals, such as those described by (2) and (4), one relies essentially on the kurtosis estimation. Two-sided tests (for \( b_2 \neq 3 \) or one-sided tests (for \( b_2 > 3 \) or \( b_2 < 3 \) can be envisaged. Consider the natural estimator of \( b_2 \), namely
\[
b_2 = m_4/m_2^2, \quad m_k = \sum(x(i) - \mu)^k/N.
\]

Asymptotically, as \( N \to \infty \), \( b_2 \) is normally distributed with mean \( E(b_2) \to 3 \) and variance \( \text{var}(b_2) \to 24/N \) [9]. Let the standardized version of \( b_2 \) be
\[
\sqrt{\beta}(b_2) = 6(N^2 - 5N + 2)/(N + 7)(N + 9) \sqrt{6(N + 3)(N + 5)/N(N - 2)(N - 3)}. \tag{5}
\]
For $N$ above about 30, $x$ has approximately a chi-square distribution with
\[ A = 6 + \frac{8}{\sqrt{\beta(b_2)}} \left[ \frac{2}{\sqrt{\beta(b_2)}} + \sqrt{1 + \frac{4}{\beta(b_2)}} \right] \] (6)
degrees of freedom (such that $A > 18$). Since $A$ is never small we can convert the chi-squared distribution to an equivalent normal distribution by the Wilson-Hilferty transformation. The Gaussian random variable corresponding to $x$ is [9]
\[ Z(b_2) = \frac{\sqrt{2}A}{\sqrt{2}} \left[ \left( 1 - \frac{2}{9A} \right) - \left( \frac{1 - (2/A)}{1 + x\sqrt{2/(A-4)}} \right) \right] \] (7)
with $Z(b_2)$ being approximately normally distributed under the hypothesis that the sequence $x(i)$ is Gaussian (hypothesis $H_0$). The decision threshold $\lambda$ can be computed by $\lambda = Q^{-1}(P_{fa}/2)$, where $P_{fa} = \Pr((Z(b_2)) > \lambda | H_0)$ denotes the probability of false alarm and the error function $Q(i)$ is
\[ Q(y) = \frac{1}{\sqrt{2\pi}} \int_{y}^{\infty} \exp(-u^2/2) \, du \] (8)

B. Lilliefors test. The Kolmogorov-Smirnov (K-S) test of goodness of fit provides a means of testing whether a set of samples belong to a completely specified continuous distribution. However, when certain parameters of the distribution are not known and have to be estimated from the samples, then the K-S test cannot be applied without modification. The Lilliefors test for normality is one of the most well-known modifications of the K-S test.

The procedure is as follows [10]: given a set of $N$ samples one determines $D = \max|F(X) - S_N(X)|$, where $S_N(X)$ is the sample cumulative distribution function and $F(X)$ is the cumulative normal distribution function with mean equal to the sample mean and variance equal to the sample variance. If $D$ exceeds a critical tabulated value, one rejects the hypothesis that the observations are from a normal distribution. The tabulated critical values were obtained by Monte Carlo calculation. For $N > 30$ the ratio of the Monte Carlo values to the standard values remains relatively constant. In that case, the resulting critical values of $D$ for different levels of significance are equal to $\gamma/\sqrt{N}$, with $\gamma$ given in [10].

Any value of $D$ which is greater than or equal to the tabulated value is significant at the indicated level of significance. If $D$ exceeds the critical value given by the table, the normality hypothesis is rejected with probability of false alarm equal to the level of significance.

C. Cramér-von Mises test. Consider the sequence $x(i)$, $i = 1, \ldots, N$, of independent identically distributed random variables with corresponding sample distribution function $F_N(x)$, $x \in \mathbb{R}$. The Cramér-von Mises statistic for testing whether the random variables have distribution function $F(x)$ (hypothesis $H_0$), where $F(x)$ is assumed to be a continuous function on $\mathbb{R}$, is [11]
\[ \omega^2 = N \int_{-\infty}^{\infty} [F_N(x) - F(x)]^2 \, dF(x) \] (9)

Consider now that $F(x)$ is the distribution function of a Gaussian r.v. with mean $\mu$ and variance $\sigma^2$. If $U_1 \leq \ldots \leq U_N$, with $U_i = F(x(i)) = 1 - Q((x(i) - \mu)/\sigma)$, then
\[ \frac{1}{12N} \leq \omega^2 = \frac{1}{12N} \sum_{i=1}^{N} \left( U_i - \frac{2i - 1}{2N} \right)^2 \leq \frac{N}{3} \] (10)
for $N = 1, 2, \ldots$. The hypothesis $H_0$ is rejected whenever $\omega^2 > \lambda$, where $\lambda$ is such that $\Pr(\omega^2 > \lambda) = \alpha$ (probability of false alarm). For $N \to \infty$, reference [12] exhibits a table of the probabilities $\Pr(\omega^2 > \lambda)$ versus the decision threshold $\lambda$.

D. Giannakis-Tsatsanis test. Consider the zero-mean, real, stationary, Gaussian sequence $x(i)$, $i = 1, \ldots, N$, with variance $\sigma^2$. The components of the 4th-order autocumulant vector are defined by [7]
\[ c_4(a,b,c) = E\{x(i)x(i+a)x(i+b)x(i+c)\} + c_2(a)c_2(b-c) - c_2(b)c_2(c-a) \] (11)
for $a > b > c \geq 0$, where $c_2(a) \equiv E\{x(i)x(i+a)\}$ is the autocorrelation function. In particular, we have $c_4(0,0,0) = E\{x^4(i)\} - 3E\{x^2(i)\}^2$ usually denoted as $\kappa_4 \ [8]$.

The classical or natural estimator of the autocumulant $c_4(a,b,c)$ is
\[ \hat{c}_4(a,b,c) = E\{x(i)x(i+a)x(i+b)x(i+c)\} - \hat{c}_2(a)\hat{c}_2(b-c) - \hat{c}_2(b)\hat{c}_2(a-c) - \hat{c}_2(c)\hat{c}_2(a-b) \] (12)
where $\hat{c}_2(a)$ is the estimate of the autocorrelation given by
\[ \hat{c}_2(a) = \frac{1}{N} \sum_{i=1}^{N-a} x(i)x(i+a) \] (13)
The Gaussianity test uses $N_c = L(L+1)(L+2)/6$ lags $(a,b,c)$, with $0 \leq c \leq b \leq a \leq (L-1)$, and $L = 1, 2, \ldots$, which are collected into an $N_c \times 1$ vector, called $\hat{c}_4$ for simplicity. Assume lexicographic sorting where the element $\hat{c}_4(p)$, with $p = 1, 2, \ldots, N_c$, is given by $\hat{c}_4(p) = \hat{c}(a,b,c)$ such that $p = (1/6)(a^3 + 3a^2 + 2a + 6) + (1/2)(b + 1) + c$.

The Gaussianity test is formulated as the following binary hypothesis testing problem: $H_0 : \hat{c}_4 \sim N(0, N^{-1}Q)$ versus $H_1 : \hat{c}_4 \sim N(c_4, N^{-1}Q)$, with $c_4 \neq 0$. Matrix $Q$ is an estimate of the asymptotic covariance matrix of $c_4$, which is $Q = \lim_{N \to \infty} N\mathbb{E}\{(\hat{c}_4 - c_4)(\hat{c}_4 - c_4)^T\}$. The test, defined as [7]
\[ t_4 = \hat{c}_4^TQ^{-1}\hat{c}_4 \] (14)
is asymptotically (when \( N \to \infty \)) chi-square distributed with \( N_c \) degrees of freedom, mean \( N_c \) and variance \( 2N_c \). Given the decision threshold \( \lambda \), the probability of false alarm is [13]

\[
P_{fa} = Pr \{ t_4 \geq \lambda | H_0 \} = \exp \left( -\frac{\lambda}{2} \right) \sum_{k=0}^{(N_c/2)-1} \frac{1}{k!} \left( \frac{\lambda}{2} \right)^k
\]

(15)

A major difficulty with the computation of \( t_4 \) is the determination of \( \hat{Q} \). For \( N \gg 1 \), this matrix is almost diagonal with the entries on the main diagonal much smaller in modulus than those of the diagonal. Therefore, we can approximate \( \hat{Q} \) by a truly diagonal matrix, i.e., \( \hat{Q} \approx \text{diag}(d_1, \ldots , d_{N_c}) \), with \( d_p = E[\hat{c}_4(p)] \), where \( \hat{c}_4(p) \) is obtained using (12).

It can be shown that the elements \( \hat{c}_4(a, b, c) \) of the cumulant vector \( \hat{c}_4 \) take no more than five different values, regardless of \( N \) and \( N_c \). We name them as types I through V according to the following table.

<table>
<thead>
<tr>
<th>Type</th>
<th>Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>((a, 0, 0)) or ((a, a, a), a &gt; 0)</td>
</tr>
<tr>
<td>II</td>
<td>((a, a, 0), a &gt; 0)</td>
</tr>
<tr>
<td>III</td>
<td>((a, a, b), (a, b, 0), a, b &gt; 0)</td>
</tr>
<tr>
<td>IV</td>
<td>((a, b, c), a &gt; b &gt; c &gt; 0)</td>
</tr>
</tbody>
</table>

The value \( d_1 \) was determined, for instance, in [14], and is given by \( d_1 = 12(2N^2 - 3N + 36)\sigma_n^6/N^3 \). For \( N \gg 1 \), which is the case considered in this paper, the dependency is \( d_1 \approx 24/N\sigma_n^6 \). By the same token, we have for the remaining four types of cumulant elements: \( d^{II} \approx (6/N)\sigma_n^8 \), \( d^{III} \approx (4/N)\sigma_n^8 \), \( d^{IV} \approx (2/N)\sigma_n^8 \), and \( d^V \approx (1/N)\sigma_n^8 \).

4. SIMULATION RESULTS

For all the simulations we consider a sinusoidal interference with frequency equal to 100 kHz or a chirp interference with frequency range \( L = 40 \) MHz, centered at the GNSS signal carrier frequency, with repetition period \( T_J = 20\mu s \). The sampling rate is \( r = 50 \) MHz.

Fig. 2 exhibits the receiver operating characteristic (ROC) for each of the four Gaussianity tests under discussion. The curves depict the pairs \((P_{fa}(\lambda), P_{fa}(\lambda))\), where \( P_{fa}(\lambda) \) stands for the probability of false alarm for the decision threshold \( \lambda \), with hypothesis \( H_0 \) and \( P_d(\lambda) \) denotes the probability of correct detection with hypothesis \( H_1 \). The number of complex observations is \( N_s = 10^4 \), which are organized as \( 2N_s \) real samples (real part of each observation followed by imaginary part), and the curves were computed with \( 5 \times 10^3 \) independent runs. In the simulations we considered that \( N_g = 10 \) independent GNSS signals with equal signal-to-carrier ratios \((C/\!\!/N_0 = 50 \) dB-Hz\) were received.

The results show identical performances of each test for the sinusoidal and chirp interference, except for the Giannakis-Tsatsanis (G-T) test, where performance achieved with the sinusoidal interference is much better than the one obtained with the chirp interference for \( N_c = 4 \) lags. However, this discrepancy is attenuated by increasing \( N_c \).

The G-T test, although more demanding in terms of computational effort (especially for a large number of lags) permits to obtain better results than the other three tests under study. Simulations have shown that the performance improves with \( N_c \) for both types of interference. This means that by increasing \( N_c \), we can improve the ability of the test to separate hypotheses \( H_0 \) and \( H_1 \) for a certain number of observations. Given the advantages over the remaining tests, we will concentrate henceforth on the analysis of the G-T test.

Consider the cost function \( C(\lambda) = |P_{fa}(\lambda) + (1 - P_d(\lambda))^{1/2} | \) given by the distance from the ROC curve to the upper left corner of the plot. The optimum value of \( \lambda \) corresponds to minimize \( C(\lambda) \). By computing the values of \( C_{\min} = \text{min}_\lambda \{C(\lambda)\} \) we can compare the performance of two tests for a given interference-to-signal power \((J/C)\), the best being the one with the smallest ROC metric \( C_{\min} \). Alternatively, we can determine the gain of test \( A \) relative to test \( B \) \((J/C)_A = (J/C)_B \) for a given value of \( C_{\min} \).

Figure 3 exhibits the curves of \( C_{\min} \) versus the ratio \( J/C \) for sinusoidal and chirp interference where different values of lags \( N_s \) were considered. In the figure, \( C_{\min} \approx P_{fa}(\lambda) \sqrt{2} \approx (1 - P_d(\lambda))^{1/2} \). The solid and the dashed curves were built with \( N_s = 10^4 \) and \( N_s = 2 \times 10^4 \) samples (observations), respectively. Note that there is an improvement of about 1 dB when the number of observations doubles. The dotted curves were obtained with the exact value of the Gaussian noise variance \( \sigma_n^2 \); for \( N_s = 35 \) and \( N_s = 10^4 \) samples (semi-blind scheme); in contrast, all the other curves result from estimating the noise variance from the observations (blind schemes). For all the curves \( C_{\min} \to 1/\sqrt{2} \) when \( J/C \to 0 \), which corresponds to the ROC curve with \( P_d = P_{fa} \). In this case, the test is unable to separate hypotheses \( H_0 \) and \( H_1 \). The curves with \( N_c = 1 \) mean that only cumulant \( c_4(0, 0, 0) \) was used to test Gaussianity, which is equivalent to the kurtosis tests, such
Simulations carried out have shown that, for practical values of the GNSS signals carrier-to-noise ratio ($C/N_0 \leq 50$ dB-Hz), the presence of the signals does not affect significantly the results. This means that about the same conclusions could be achieved if the signals without interference in (2) or (4) were solely white Gaussian complex sequences $N_m$.

5. CONCLUSION

This paper aimed at analyzing the performance of several Gaussianity tests as a blind method to detect narrow (sinusoidal) and wideband (chirp) interference in GNSS signals. The tests under analysis can be classified into two classes: the ones that resort to the computation of moments or cumulants (Anscombe-Glynn and Giannakis-Tsatsanis), and those that rely on the divergence of the empirical distribution function relative to the theoretical Gaussian distribution (Lilliefors and Cramér-von Mises). Simulations have shown that the Giannakis-Tsatsanis test produces the best results at the cost of a higher computational burden. In general, this test is more sensitive to narrowband interference, thus meaning that more processing effort is required to detect chirp interference. The test can be used as a benchmark for comparison with other interference detection techniques, such as those based on the eigenvalues of the sample covariance matrix [15].

REFERENCES